



Research to Practice Brief

Welcome to the fourth University of Minnesota College Readiness Consortium and Educational Equity Resource Center's *Research to Practice Brief*. *Research to Practice Briefs* provide school leaders with information and resources on current Minnesota issues related to closing achievement gaps and preparing students for postsecondary success. School leaders can use the *Research to Practice Brief* as a framework for applying current research to best meet their school's needs.

Issue Four: Measuring and Reporting Student Achievement

2015 Legislation

Educators are increasingly expected to understand, use and report data on students from various assessments that serve different purposes. Yet assessments change regularly, such as periodic revisions of the Minnesota Comprehensive Assessments (MCAs), and new ways of calculating progress are adopted, such as the Multiple Measurements Rating. In 2015 the Minnesota legislature modified the graduation assessments requirement, for example, students are no longer required to take the ACT as a college readiness assessment. Nationally the recent 2015 federal reauthorization of ESEA (commonly known as NCLB) into the 'Every Student Succeeds Act' will lead to further changes. Education leaders with measurement and assessment skills will be prepared to meet the demands of the changing assessment landscapes.

Principles from the Research

We know assessment matters. Teachers assess students every day by questioning and listening to student comments in class,



evaluating projects, reviewing homework and yes, creating and correcting tests. Assessments impact students in valuable ways by communicating learning goals, including the subject-matter content and thinking processes prioritized by the teacher. Students typically get the clearest and most explicit message about which content and skills are most important in the field by what shows up on the test.

If we don't develop, use and interpret appropriate assessments, as well as provide an instructional response based on the data generated, students receive messages that are inaccurate or unhelpful to their academic growth. Also, measures of school quality can be based on factors that are correlated with student characteristics rather than the quality of its program. In order to more accurately and effectively use assessments to increase student achievement, three principles of measurement are particularly relevant:

- ❖ **All Measurements have Error:** All tests include some degree of error, because of the way we interpret and use test scores. We typically want test scores to tell us what students know and can do on broader content than what was included on the test, in contexts that were not tested, under novel conditions and beyond the test day. The need to generalize over content, contexts, conditions and time means that the single test score is less precise—since we've only sampled some content or contexts, under one condition at a single point in time. We can interpret scores with the standard error of measurement reported in technical manuals, but this includes only one source of test score error (the sample of items).

- ❖ **School accountability tests are not diagnostic:** Large-scale, summative tests designed to assess schools do not provide the type of fine-grained diagnostic information teachers need to adjust their instruction to enhance learning. For example, the strands on the MCAs are not diagnostic, and should not be used alone to determine interventions, because the number of items is too small to deeply cover content studied over a long period of time and do not provide unique information above and beyond the total score. Accountability tests may provide initial indication of strengths and weaknesses for groups, to be confirmed and understood with additional information.
- ❖ **Multiple measures are required:** One measure is insufficient: Test scores are point-in-time estimates rather than exact measures. Tests have only a sample of all possible questions, and students will fare differently with different questions. Motivation, distractions, and other factors can influence results. Multiple measures greatly increase the accuracy of the overall assessment of a student.

These principles are interrelated. School leaders can be champions of their message and improve the use of test score results dramatically. They remind us that test performance for any student is temporal, and to maximize the use of data to improve teaching and learning, multiple sources of information is required. Another implication is that not every piece of information is needed on every student every year. For example, the Minnesota Student Survey is conducted every three years, and now contains information regarding social-emotional learning important for success in school and beyond. Such information is relatively stable over time and data collection should be collected as it is needed to inform continuous improvement.

How Should School Leaders Apply the Research?



Local educators cannot change state assessments; they can affect interpretation and use. How test scores are reported and used are equally daunting challenges but are within an administrator's control and responsibility. Not every education leader has received sufficient training on measurement and assessment, so it is important to take steps to learn and develop a basic understanding. Such understanding can be supported through the assessment expertise available among district staff, where it exists, or in other resources, including a recent guidance document developed in collaboration with the Minnesota Assessment Group (see resources below).

First, effective use of data requires attention to the specific context being assessed. High mobility and/or dropout rates, or small group sizes, can distort results for the total student population and signal the need for more information. The five racial/ethnic groups used in the Minnesota reporting systems may not be representative of the demographics actually present in a given school or the heterogeneity within racial/ethnic groups. Similarly, income is continuous, yet students are categorized into two groups – low or high income (often based on Free/Reduced lunch status, failing to consider other assets including social or cultural capital).

Second, despite the caveats just stated it is important to disaggregate data to look at relevant groups of students. When analyzing or reporting on gaps between groups, districts should be explicit about how the groups are defined and the impact of their local context. The more we know about our students, the more meaningful such information will be.

Third, leaders should use multiple measures (such as different exams, grades, observations, etc.) and be prepared to explain their purpose and the reasons for choosing them. If, for example, MAP and MCA results produce similar interpretations (as one may mimic the other), reporting multiple sources provides additional

support for conclusions. If results differ, explore their different contents. In many cases this may require innovative assessments of what students know and can do, recognizing the unique contexts students bring.

Fourth, consider multiple methods of presenting the same data. It is far too easy to present an average score, or percent proficient, for a group and leave the impression that it adequately describes all students in the group. Students and their families internalize unintended messages. Instead, consider presenting the range of scores for students as it may show that each group includes students who score high and low. Consider a longitudinal analysis monitoring the achievement of the same group of students over time, from one grade to the next.

Finally, school leaders should discuss with their staff and each other about the role assessment serves. Some assessments are system checks and ought to be identified as such. Others, such as common assessments that teachers write together, can be used to determine students' strengths *and* areas of need but also to help teachers come to agreement on instructional goals (and potentially deeper understanding of the standards). The challenge for schools is to create and communicate a balanced assessment system using the strengths of assessments to address learning, instructional, and accountability needs.

What Does Effective Data Use Look Like?



An effective school leader accepts responsibility for accuracy and transparency of all data, including assessments. The school adopts principles for appropriate and fair data use such as:

1. Data will be collected for clearly stated purposes and uses, for which validity evidence exists. Validity is the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests.

2. Principles of efficiency, effectiveness, and equity will drive improvement efforts regarding data collection, reporting, and use. Data will be employed in a positive manner for continuous improvement, not as part of a negative campaign or for blaming.
3. Data that go unused will be reviewed for appropriateness, meaningfulness, and usefulness, and data collection tools will be modified if possible or removed from use.
4. Collection and reporting strategies will be periodically reviewed, on the above principles.

In all presentations of assessment data, school leaders are able to place results in context that makes data actionable to enhance the efficiency, effectiveness, and equity of the school's efforts. Data reporting acknowledges variability within groups and relevant differences between groups – with the goal of understanding and meeting the needs of all students. School leaders demand evidence that supports the intended interpretations and uses of all data, particularly test scores. Where evidence is not forthcoming, leaders limit use of test scores to only those uses that have been validated, and are pragmatically useful to the learner, educators, families, and the learning system.

Resources

Professional Development

The University of Minnesota houses one of the largest [Educational Measurement Programs](#) in the country, offering many options for additional coursework, seminars, and PD opportunities.

[*Competency Standards in Assessment for Educational Administrators*](#)

[*Classroom Assessment Standards for PreK-12 Teachers*](#)

Practical Reports on Test Use

Analyzing & Reporting Achievement Gaps: Guidance for Minnesota Schools

Lessons Learned about Testing, a frank and expert summary of practical lessons on test use.

Other Resources

The University of Minnesota's Center for Applied Research and Educational Improvement (CAREI) conducts research and evaluations on education topics and provides assessment support to educators and schools. Browse their publications and explore their services online.

The Buros Center for Testing provides resources for improving Assessment Literacy.

The National Council on Measurement in Education provides briefs on Instructional Topics in Ed Measurement.



The Research to Practice Brief is created by the U of M's College Readiness Consortium and the Educational Equity Resource Center. Many thanks to Michael Rodriguez, who contributed ideas and research for this edition. For more information about current education research, read CEHD's Vision 2020 blog: <http://cehdvision2020.umn.edu/cehd-blog/>. For questions about the Brief, contact Julie Sweitzer, Director, College Readiness Consortium and Co-Director, Educational Equity Resource Center, sweitzer@umn.edu.

Copyright © 2016 Regents of the University of Minnesota, All rights reserved.